J. Mol. Biol. (1997) **274**, 303–309

JMB

COMMUNICATION

DNA-binding Specificity of the Homeodomain-leucine Zipper Domain

Giovanna Sessa[1], Giorgio Morelli[2] and Ida Ruberti[1]*

[1]*Centro di studio per gli Acidi Nucleici, c/o Dipartimento di Genetica e Biologia Molecolare Università di Roma La Sapienza, P. le Aldo Moro 5 00185 Rome, Italy*

[2]*Unità di Nutrizione Sperimentale, Istituto Nazionale della Nutrizione, Via Ardeatina 546, 00178 Rome, Italy*

**Corresponding author*

Homeodomain-leucine zipper (HD-Zip) proteins are putative transcription factors identified only in plants. The study of the DNA-binding properties of the ATHB-1 and -2 HD-Zip (HD-Zip-1 and -2) domains showed that they interact with DNA as homodimers and recognize two distinct 9 bp pseudopalindromic sequences, CAAT(A/T)ATTG (BS-1) and CAAT(G/C)ATTG (BS-2), respectively, as determined by selecting high-affinity binding sites from random-sequence DNA. Here, we report a mutational analysis of the HD-Zip-2 domain. We determined that conserved amino acid residues of helix 3, Val47 and Asn51, and Arg55 are essential for the DNA-binding activity of the HD-Zip-2 domain. We demonstrated that the preferential recognition of a G/C base-pair at the central position by the HD-Zip-2 domain is abolished either by the replacement of Arg55 with lysine or by the substitution of Glu46 and Thr56 with the corresponding residues of the HD-Zip-1 domain (alanine and tryptophan, respectively). In contrast, substitution of Arg55 with lysine in the HD-Zip-1 domain significantly reduced DNA-binding activity without changing the specificity of recognition. Finally, we determined that differences in residues outside helix 3 further contribute to the DNA-binding specificity of the HD-Zip domain. Taken together, the data strongly suggest that the preferential recognition of BS-2 and -1 by the HD-Zip-2 and -1 domains, respectively, may be attributable to a distinct orientation of the side-chain of Arg55 in these two domains.

Keywords: Arabidopsis; HD-Zip domain; homeodomain; leucine zipper; protein–DNA complexes

The homeodomain-leucine zipper (HD-Zip) factors define a class of HD proteins that seems to be peculiar to plants (Ruberti *et al.*, 1991; Mattsson *et al.*, 1992; Schena & Davis,1992). On the basis of sequence homologies, HD-Zip factors have been grouped into four families, HD-ZIP I to IV (Sessa *et al.*, 1994), and the complete alignment of the HD-Zip proteins identified so far has been published by Hoge and co-workers (Meijer *et al.*, 1997).

DNA-binding studies of ATHB-1 and -2, members of the *Arabidopsis* HD-ZIP I and II protein families, respectively, indicated that HD-Zip proteins interact with DNA recognition elements in a fundamentally different fashion from the classic HD proteins (Sessa *et al.*, 1993). DNA-binding analysis performed using random-sequence DNA templates showed that the ATHB-1 HD-Zip (HD-Zip-1), but not the ATHB-1 HD alone, binds to DNA. The HD-Zip-1 domain recognizes a 9 bp pseudopalindromic sequence (CAAT(A/T)ATTG, BS-1), as determined by selecting high-affinity binding sites from random-sequence DNA. Gel retardation assays performed using ATHB-1 derivatives with different lengths proved that the HD-Zip-1 domain binds DNA as a dimer. Moreover, the analysis of the DNA-binding activity of ATHB-1 derivatives indicated that a correct spatial relationship between the HD and the Zip is essential for DNA binding. We obtained similar results with the ATHB-2 HD-Zip domain (HD-Zip-2), and determined that it recognizes a distinct 9 bp pseudopalindromic sequence (CAAT(G/C)ATTG, BS-2; Sessa *et al.*, 1993).

Comparison of the HD-Zip–DNA complexes with the Antennapedia (Otting *et al.*, 1990), Engrailed (Kissinger *et al.*, 1990) and MATα2 (Wolberger *et al.*, 1991) HD–DNA complexes, and the GCN4 b-Zip–AP-1 complex (Ellenberger *et al.*, 1992), suggested that the angle between the recognition helix and the dimerization domain is larger

0022–2836/97/480303–07 $25.00/0/mb971408

Table 1. Amino acid sequences of HD-Zip-2, -1, and substitution variants

Peptide sequence																		
HD-Zip-2	42-	A	R	Q	V	E	V	W	F	Q	N	R	R	A	R	T	K	-57
HD-Zip-2 V47A	42-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-57
HD-Zip-2 Q50A	42-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-57
HD-Zip-2 N51A	42-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-57
HD-Zip-2 R55A	42-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-57
HD-Zip-2 R55K	42-	-	-	-	-	-	-	-	-	-	-	-	-	K	-	-	-	-57
HD-Zip-2 A42P	42-	P	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-57
HD-Zip-2 E46A	42-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-57
HD-Zip-2 T56W	42-	-	-	-	-	-	-	-	-	-	-	-	-	-	W	-	-	-57
HD-Zip-2 A42P E46A	42-	P	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-57
HD-Zip-2 A42P T56W	42-	P	-	-	-	-	-	-	-	-	-	-	-	-	W	-	-	-57
HD-Zip-2 E46A T56W	42-	-	-	-	-	A	-	-	-	-	-	-	-	-	W	-	-	-57
HD-Zip-2 A42P E46A T56W	42-	P	-	-	-	A	-	-	-	-	-	-	-	-	W	-	-	-57
HD-Zip-1	42-	P	-	-	-	A	-	-	-	-	-	-	-	-	W	-	-	-57
HD-Zip-1 R55K	42-	P	-	-	-	A	-	-	-	-	-	-	-	K	W	-	-	-57

The DNA fragments corresponding to the HD-Zip-2 and -1 domains were obtained as described (Sessa *et al.*, 1993; Di Cristina *et al.*, 1996). The DNA fragments coding for the substitution variants were obtained by PCR reactions as described by Sarker & Sommer (1990) and cloned into pGEM-4Z as described by Sessa *et al.* (1993). *In vitro* transcription and translation was performed as described (Sessa *et al.*, 1993). The amounts of *in vitro*-translated peptides were monitored by SDS-PAGE analysis as described by Sessa *et al.* (1993). Replacement of alanine at position 47, 50, 51 or 55 of the HD-Zip-2 domain was obtained by PCR amplification of the *ATHB-2* cDNA (Carabelli *et al.*, 1993) with ATHB-2 V47A (AGACAAGTGGAAGCTTGGTTTCAG, nucleotides 1230 to 1253 of the *ATHB-2* genomic sequence; the underlined triplet corresponds to the substitution), ATHB-2 Q50A (GAAGTTTGGTTTGCTAACA-GACGAGC, 1239 to 1264), ATHB-2 N51A (TGGTTTCAGGCTAGACGAGCA, 1245 to 1265), ATHB-2 R55A (AGACGAGCAGCTACAAAGCT, 1257 to 1267/1356 to 1364), in combination with ATHB-2 B3D (Sessa *et al.*, 1993), respectively. These PCR products were used as primers in combination with ATHB-2 A5D (Sessa *et al.*, 1993) to generate the DNA fragments corresponding to HD-Zip-2 V47A, HD-Zip-2 Q50A, HD-Zip-2 N51A and HD-Zip-2 R55A, respectively. Substitution of lysine at position 55 of the HD-Zip-2 domain was obtained by PCR reaction with primers ATHB-2 R55K (CAGACGAGCAAAAA-CAAAGCTG, 1256 to 1267/1356 to 1365) and ATHB-2 B3D; the PCR product was then used in combination with ATHB-2 A5D. Replacement of proline at position 42, alanine at position 46, and tryptophan at position 56 of the HD-Zip-2 domain was obtained by PCR reactions with ATHB-2 A42P (AGGGTTAC-GACCAAGACAAGTG, 1217 to 1238), ATHB-2 E46A (AGACAAGTGGCTGTTTGGTTTC, 1230 to 1251), ATHB-2 T56W (CGAGCAAGATGGAAGCTGA, 1260 to 1267/1356 to 1366), in combination with ATHB-2 B3D, respectively. These PCR products were used as primers in combination with ATHB-2 A5D to generate the DNA fragments corresponding to HD-Zip-2 A42P, HD-Zip-2 E46A and HD-Zip-2 T56W, respectively. To construct the double mutants, the DNA fragments corresponding to HD-Zip-2 E46A and HD-Zip-2 T56W were used as templates in PCR reactions with either ATHB-2 A42P or ATHB-2 E46A in combination with ATHB-2 B3D, respectively. These PCR products were used as primers in combination with ATHB-2 A5D. The fragment corresponding to HD-Zip-2 A42P E46A T56W was obtained by PCR amplification of the fragment corresponding to HD-Zip-2 E46A T56W with ATHB-2 A42P and ATHB-2 B3D; the PCR product was then used as a primer in combination with ATHB-2 A5D. Replacement of lysine at position 55 of the HD-Zip-1 domain was obtained by PCR amplification of the *ATHB-1* cDNA (Ruberti *et al.*, 1991) with ATHB-1 R55K (CGCCGAGCTAAGTGGAAAACA nucleotides 676 to 696 of the *ATHB-1* cDNA) and ATHB-1 3D (Sessa *et al.*, 1993). This PCR product was used as a primer in combination with ATHB-1 5D (Sessa *et al.*, 1993) to generate the DNA fragment corresponding to HD-Zip-1 R55K.

with respect to that observed in the GCN4–DNA complex (Sessa *et al.*, 1993). By superimposing the recognition helices such that Asn51 of the HD-Zip domain and Ala239 of GCN4 are located in the same spatial orientation, we deduced the putative contacts in the HD-Zip–DNA complex (Sessa *et al.*, 1993). We proposed that conserved amino acid residues of helix 3 (Val47, Gln50 and Asn51) that make specific base contacts in HD–DNA complexes (Otting *et al.*, 1990; Kissinger *et al.*, 1990; Wolberger *et al.*, 1991) make base contacts in the HD-Zip–DNA complex. Moreover, in analogy with the role of Arg243 in the GCN4 b-Zip–AP-1 complex (Ellenberger *et al.*, 1992), we suggested that Arg55 of one of the two HD-Zip monomers is responsible for recognition of the central position of the pseudopalindromic site in the HD-Zip–DNA complex; Arg55 of the other monomer may form hydrogen bonds to the phosphate backbone. It is worth mentioning that the asymmetric GCN4 contacts at the center of the DNA-binding site have been observed only in the structure of the GCN4

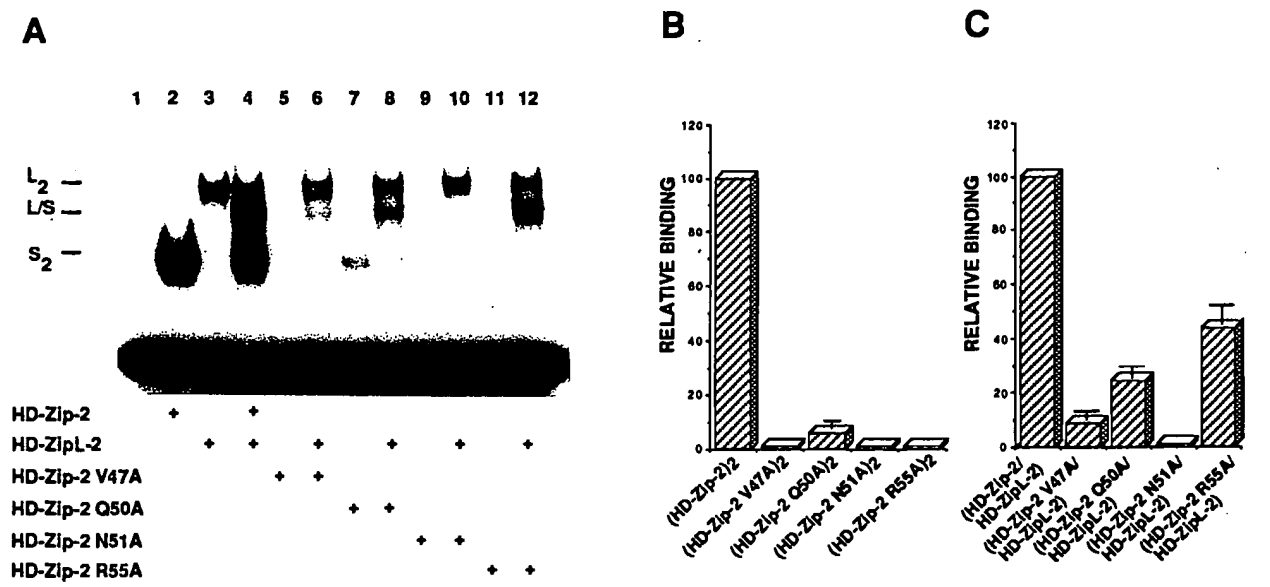

Figure 1. Amino acid residues at conserved positions of the HD-Zip domain play a major role in DNA recognition. a, DNA-binding activities of HD-Zip-2 and its variants. Equal amounts of the indicated proteins in the short form (S) were assayed for binding to the BS-2 probe individually (lanes 2, 5, 7, 9 and 11) or in combination with the HD-Zip-2 peptide in a longer form (HD-ZipL-2, lane 3; Di Cristina *et al.* (1996); lanes 4, 6, 8, 10 and 12). Note that mixed-dimer formation is observed as an additional band (L/S) with intermediate mobility (lanes 4, 6, 8 and 12). The control (lacking added RNA) is shown in lane 1. Gel retardation assays were performed as described (Di Cristina *et al.*, 1996). b and c, Quantification of DNA-binding activities of HD-Zip-2 variants expressed as relative binding, which refers to the DNA-binding activity of the HD-Zip-2 domain in the S_2 form in b and in the S/L form in c. Relative amounts of protein–DNA complexes were determined by using an Imaging densitometer GS-670 (Bio-Rad).

b-Zip domain in complex with the pseudopalindromic AP-1 site. In contrast, the GCN4 contacts to the central nucleotides of the palindromic ATF/CREB sequence are symmetric, with Arg243 of each subunit bridging between the central G of its half-site and a neighboring DNA phosphate group (Konig & Richmond, 1993; Keller *et al.*, 1995).

The hypothesis that Arg55 of one of the two HD-Zip monomers may form base contacts in the HD-Zip–DNA complex has been recently supported by the finding that Arg124 (corresponding to Arg55 of the a1 HD) of the a1 protein in the a1/α2 homeodomain heterodimer–DNA complex is involved in base contacts (Li *et al.*, 1995). Arg55 of the a1 HD forms two hydrogen bonds with a guanine base, and donates an additional hydrogen bond to the thymine base that is base-paired with the adenine base that is contacted by the invariant Asn51 (Li *et al.*, 1995).

To evaluate the functional relevance of our predictions, the DNA–protein interactions were investigated by mutagenesis of the HD-Zip-2 domain. Val47, Gln50, Asn51 and Arg55 were substituted with alanine (Table 1). The presence of alanine at position 47, 51 or 55 prevents the HD-Zip-2 domain from binding, whereas at position 50 it significantly reduces HD-Zip-2 DNA binding activity (Figure 1a and b). To further investigate the role of Val47, Asn51 and Arg55, we analyzed by gel retardation assay the formation of mixed dimers between a long derivative of ATHB-2 and a short one containing the mutation. The presence of alanine at either position 47 or 51 in only one of the two subunits strongly affected the ability of the HD-Zip-2 domain to bind to DNA, whereas the replacement of arginine with alanine at position 55 of one of the two subunits affected DNA binding only slightly (Figure 1a and c). This is consistent with the involvement of Arg55 in the recognition of the central base-pair of the BS-2 pseudopalindromic site that should be alternatively recognized by one of the two subunits.

The analysis of the DNA-binding properties of GCN4 showed that the peptide binds strongly to the asymmetric AP-1 site (ATGAGTCAT) and, also, with a reduced relative affinity, to a target variant with a central A/T pair (ATGAATCAT: Suckow *et al.*, 1994). Thus, we analyzed the ability of the HD-Zip-2 domain to recognize a site in which the central base-pair was an A/T pair (BS-1). Gel retardation assays demonstrated that the HD-Zip-2 domain, although less efficiently, binds to BS-1 (Figure 2a). This suggested that Arg55 in the context of the HD-Zip-2 domain is also able to recognize an A/T base-pair at the central position. The preferential binding of the HD-Zip-2 domain to a G/C pair at the central position may be attributed to the ability of the guanidinium group of Arg55 to establish two hydrogen bonds to O6/N7 of the central guanine base, and only one to N7 of the adenine or to O4 of the thymine base. Therefore, the role of Arg55 in the HD-Zip-2 domain

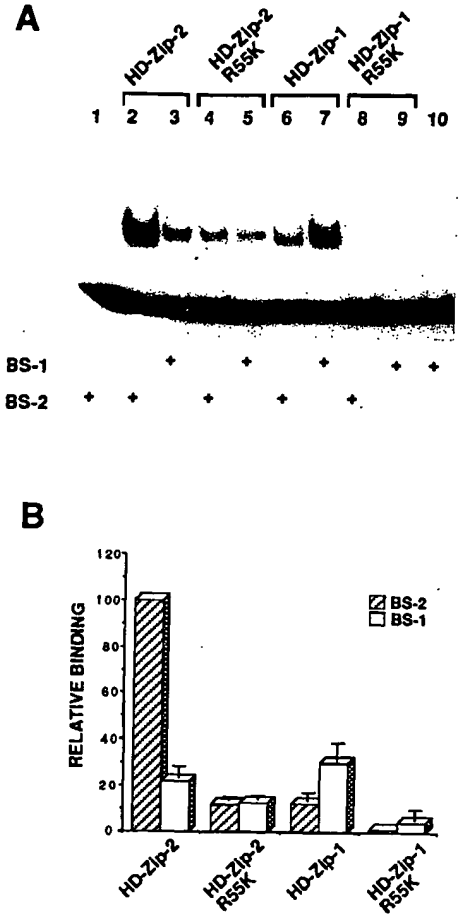

Figure 2. DNA-binding specificity of the HD-Zip-2 and 1 domains. a, DNA-binding activities of the HD-Zip-2 and -1 peptides and their variants. Equal amounts of the indicated proteins were assayed for binding to BS-2 (lanes 2, 4, 6 and 8) and BS-1 (lanes 3, 5, 7 and 9). The controls are shown in lanes 1 and 10. b, Quantification of DNA-binding activities of the peptides shown in a expressed as relative binding, which refers to the DNA-binding activity of the HD-Zip-2 domain.

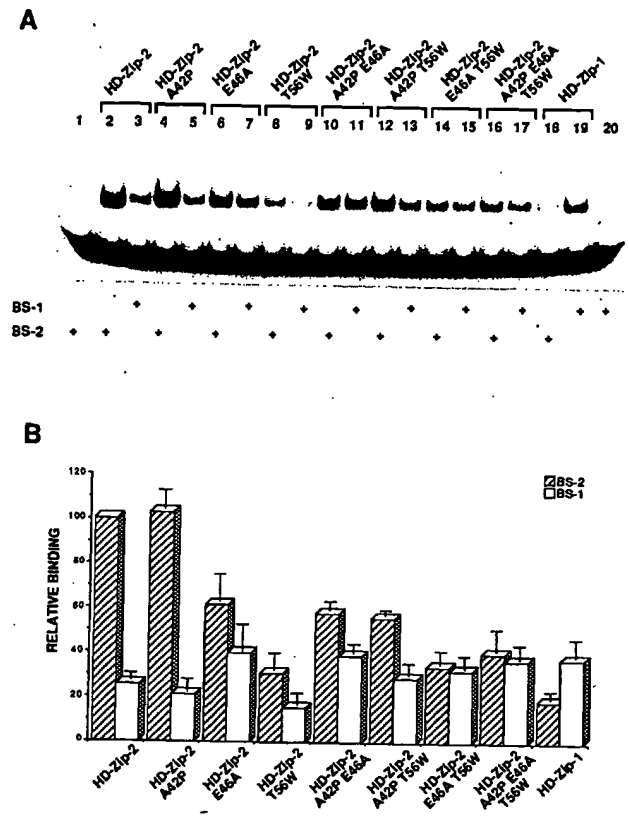

Figure 3. Amino acid residues at positions 46 and 56 influence DNA-binding specificity of the HD-Zip domain. a, DNA-binding activities of HD-Zip-2, its substitution variants, and HD-Zip-1. Equal amounts of the indicated proteins were assayed for binding to BS-2 (lanes 2, 4, 6, 8, 10, 12, 14, 16 and 18) and BS-1 (lanes 3, 5, 7, 9, 11, 13, 15, 17 and 19). The controls are shown in lanes 1 and 20. b, Quantification of DNA-binding activities of the peptides shown in a expressed as relative binding, which refers to the DNA-binding activity of the HD-Zip-2 domain.

was further investigated by substituting the arginine with lysine (Table 1), which could form only one hydrogen bond in either case. Consistently, the replacement of Arg55 with lysine in the HD-Zip-2 domain had essentially no effect on the ability of the peptide to recognize an A/T base-pair at the central position, whereas it reduced the recognition of a G/C pair at the central position (Figure 2a). From these data (Figure 2b) we conclude that a G/C pair is recognized by the mutant HD-Zip-2 domain essentially as the A/T pair. In the HD-Zip-1 domain the contact with either adenine or thymine at the central position, although energetically less favored, is preferred to that with a guanine base (Figure 2a and b), suggesting that the side-chain of the arginine residue at position 55 in this domain has a distinct spatial orientation. The substitution of Arg55 with lysine (see Table 1) reduced DNA binding but did not change the specificity of the recognition (Figure 2a and b). It is likely that in the context of the HD-Zip-1 domain, the smaller side-chain of lysine is less effective than arginine in making a hydrogen bond with adenine or thymine and, to an even greater extent, with guanine.

To support our hypothesis, we investigated whether amino acid residues of helix 3 of the HD-Zip-1 domain (amino acid residues 42 to 57; Ruberti *et al.*, 1991; Sessa *et al.*, 1993) are able to influence the preferential recognition of a G/C base-pair at the central position of the binding site by the HD-Zip-2 domain. To this end, we constructed a series of mutant HD-Zip-2 peptides. In particular, the amino acid residues present at positions 42, 46 and 56 in the HD-Zip-2 domain were replaced with those of the HD-Zip-1 domain (Table 1). The analysis of the DNA-binding specificity of the mutant peptides indicated that the substitution of Ala42 with proline has essentially no effect, whereas the replacement of Glu46 with alanine as well as Thr56 with trypto-

phan affects the ability of the HD-Zip-2 domain to distinguish between a G/C and an A/T base-pair at the central position (Figure 3a and b). The double mutant with alanine in position 46 and tryptophan in position 56 binds to the BS-1 site with a relative affinity comparable to that of the wild-type peptide, and with a significantly reduced affinity to the BS-2 site (Figure 3a and b). The triple mutant, which contains the entire helix 3 of the HD-Zip-1 domain, binds to the BS-2 and BS-1 sites with a relative affinity comparable to that of the double mutant. From these data we conclude that the residues 46 and 56 of helix 3 affect the DNA-binding specificity of the HD-Zip domain. It is worth mentioning that the amino acid residues at positions 46 and 56 are Ala and Trp in all members of the *Arabidopsis* HD-ZIP I family, and Glu and Thr in all proteins of the *Arabidopsis* HD-ZIP II family, respectively (see Meijer *et al.*, 1997 and references therein).

The analysis of the DNA-binding properties of HD-Zip-2, -1, and mutant HD-Zip-2 peptides (Figure 3) also suggested that differences in regions outside helix 3 may indirectly affect the DNA-binding specificity of the HD-Zip domain. In the attempt to identify some additional contribution to the DNA-binding specificity of the HD-Zip domain, we investigated whether the amino acid residues 58 to 67 of the HD-Zip-1 domain (corresponding to the residues of the fork region in the b-Zip domain, Ellenberger *et al.*, 1992) are able, in combination with helix 3 of the HD-Zip-1 domain (amino acid residues 42 to 57), to affect the specificity of the HD-Zip-2 domain such that the recognition of an A/T base-pair at the central position of the binding site is preferred to that of a G/C pair. To this end, we constructed two mutant HD-Zip-2 peptides in which amino acid residues 42 to 61 or 42 to 67 were replaced with those of the HD-Zip-1 domain (peptides III and IV, respectively; Figure 4a), and compared the specificity of these mutant peptides with those of HD-Zip-2 (peptide I, Figure 4a), HD-Zip-1 (peptide VI, Figure 4a), and the chimeric HD-Zip-2 peptide containing helix 3 of the HD-Zip-1 domain (peptide II, Figure 4a). The peptides III and IV bind to both BS-2 and BS-1 sites with a relative affinity comparable to that of the peptide II (Figure 4b and c). This suggested that differences in residues 1 to 41 of the HD-Zip domain may contribute to the DNA-binding specificity, probably affecting the docking of helix 3 in the major groove of the recognition site. Consistently, the replacement of helix 1 and helix 2 of peptide IV with those of the HD-Zip-1 domain (peptide V, Figure 4a) had essentially no effect on the ability of the peptide to recognise an A/T base-pair, whereas it reduced the recognition of a G/C pair at the central position (Figure 4b and c).

Taken together, the data strongly support our model of DNA binding by the HD-Zip proteins (Sessa *et al.*, 1993). First, we found that in both HD-Zip-1 and -2 domains the arginine residue at position 55 is required for DNA binding. Second, we showed that the preferential recognition of a G/C base-pair *versus* an A/T pair at the central position by the HD-Zip-2 domain is abolished either by the replacement of Arg55 with lysine or by the substitution of Glu46 and Thr56 with the corresponding amino acid residues of the HD-Zip-1 domain. Third, we found that differences in residues outside helix 3 contribute to the DNA-binding specificity of the HD-Zip domain. The data strongly suggest that the preferential recognition of BS-2 and -1 by the HD-Zip-2 and -1 domains, respectively, may be attributable to a distinct conformation of the Arg55 side-chain involved in base contacts at the center of the DNA-binding site in the HD-Zip-2 and HD-Zip-1–DNA complexes. Finally, we proved that the amino acid residues at positions 47, 50 and 51 are essential for the DNA-binding activity of the HD-Zip domain, as observed for the homeodomain (for a review, see Gehring *et al.*, 1994). Despite this, the HD-Zip domain is clearly different from the classic homeodomain, which typically binds DNA as a monomer. In the homeodomain, helix 3 is precisely oriented and stabilized for binding in the major groove by its interactions with helices 1 and 2, and by binding of the HD arm in the adjacent minor groove (Gehring *et al.*, 1994). The HD-Zip domain might have lost the ability to make base contacts in the minor groove of the DNA to prevent unfavorable steric interactions between the two N-terminal arms, and as a consequence of this its overall DNA-binding energy might be too low to bind DNA in the monomeric form. Alternatively, it might be speculated that both the classic HD and the HD-Zip domain could have evolved from a primordial homeodomain with a modest affinity for DNA. Further DNA contacts formed with bases in the minor groove and complex formation with a second protein could have evolved as independent mechanisms to augment DNA-binding specificity and affinity of such a precursor. The mechanism of binding of the yeast a1 HD supports this view. It does not bind DNA detectably as a monomer; however, it forms a complex with α2, which binds DNA with high specificity and affinity. Furthermore, the N-terminal arm of a1, unlike that of the α2 homeodomain, does not contact DNA in the a1/α2 HD heterodimer–DNA complex (Li *et al.*, 1995). In this context, it is also worth mentioning the binding mechanism of the Skn-1 protein. Skn-1 contains a basic region similar to those of b-Zip proteins but it lacks a leucine zipper dimerization domain. The Skn-1 basic region lies at the carboxyl terminus of an 85 residue domain that binds to a b-Zip half-site and also recognizes adjacent 5′ A+T-rich sequences in the minor groove through an N-terminal arm related to those of HD proteins (Blackwell *et al.*, 1994). Thus, evolution appears to have been capable of deriving different solutions to the problem of how to raise DNA binding affinity. The close linking of a leucine zipper to an HD might be an independent way that has evolved to

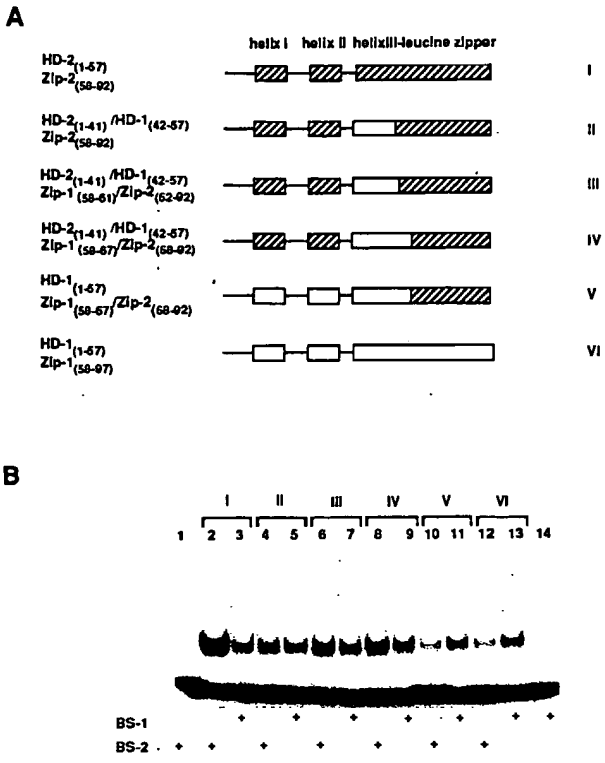

Figure 4. Amino acid residues outside helix 3 influence DNA-binding specificity of the HD-Zip domain. a, Representation of HD-Zip-2, -1 and chimeric peptides. The DNA fragments corresponding to HD-$2_{(1-57)}$/Zip-$2_{(58-92)}$ (I), HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$2_{(58-92)}$ (II), and HD-$1_{(1-57)}$/Zip-$1_{(58-97)}$ (VI) were obtained as described for Table 1 (indicated as HD-Zip-2, HD-Zip-2 A42P E46A T56W and HD-Zip-1, respectively). The DNA fragments coding for HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-61)}$/Zip-$2_{(62-92)}$ (III), HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-67)}$/Zip-$2_{(68-92)}$ (IV), and HD-$1_{(1-57)}$/Zip-$1_{(58-67)}$/Zip-$2_{(68-92)}$ (V) were obtained by PCR reactions as described by Horton *et al.* (1989). To generate the DNA fragments corresponding to peptides III and IV we first generated a DNA fragment corresponding to HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-97)}$. The fragment corresponding to HD-$2_{(1-41)}$ was obtained using the oligonucleotide ATHB-2 A5D (Di Cristina *et al.*, 1996) in combination with ATHB-2 F3D (TGCCTTGGTCGTAACCCTAATTGTTTAGCC, nucleotides 1199 to 1220 of the *ATHB-2* genomic sequence; the underlined sequence corresponds to nucleotides 646 to 653 of the *ATHB-1* cDNA); the fragment corresponding to HD-$1_{(42-57)}$/Zip-$1_{(58-97)}$ was obtained using ATHB-1 A5D (AGGGTTACGACCAAGGCAAGTGGCTGTC, 1217 to 1226; 646 to 663) and ATHB-1 3D (Sessa *et al.*, 1993). The primers ATHB-2 F3D and ATHB-1 A5D were designed so that the ends of the two PCR products contained complementary sequences. The PCR products were mixed and amplified to generate the recombinant fragment. The fragments corresponding to HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-61)}$ and HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-67)}$ were obtained by PCR amplification of the DNA corresponding to HD-$2_{(1-41)}$/HD-$1_{(42-57)}$/Zip-$1_{(58-97)}$ with the oligonucleotide ATHB-2 A5D in combination with either ATHB-1 3III (CGCAGTCTACCTCAAGCTGTTTTGTTTTCCA, 1375 to 1387; 688 to 705) or ATHB-1 3IV (GCATCTCCGGAGAAGATCGTAGTCTCT, 1395 to 1403; 709 to 726); the fragments corresponding to the different portions of the ATHB-2 leucine zipper (Zip-$2_{(62-92)}$ and Zip-$2_{(68-92)}$) were obtained using ATHB-2 5III (AAACAGCTTGAGGTAGACTGCGAGTTC, 1375 to 1392; 697 to 705) or ATHB-2 5IV (ACGATCTTCTCCGGAGATGCTGCGAGAAT, 1395 to 1412; 716 to 726) in combination with ATHB-2 B3D (Di Cristina *et al.*, 1996). The primers ATHB-1 3III and ATHB-2 5III, and ATHB-1 3IV and ATHB-2 5IV were designed so that the ends of the two PCR products contained complementary sequences. These PCR products were mixed and amplified to generate the recombinant fragments corresponding to peptides III and IV. The DNA fragment corresponding to HD-$1_{(1-57)}$/Zip-$1_{(58-67)}$ was obtained by PCR amplification of the *ATHB-1* cDNA with ATHB-1 5D (Sessa *et al.*, 1993) and ATHB-1 3IV; the fragment corresponding to Zip-$2_{(68-92)}$ was obtained using the oligonucleotides ATHB-2 5IV and ATHB-2 B3D. The two PCR products were mixed and amplified to generate the recombinant fragment corresponding to peptide V. b, DNA-binding activities of HD-Zip-2, -1 and chimeric peptides. Equal amounts of the indicated proteins were assayed for binding to BS-2 (lanes 2, 4, 6, 8, 10 and 12) and BS-1 (lanes 3, 5, 7, 9 and 11). The controls are shown in lanes 1 and 14. c, Quantification of DNA-binding activities of the peptides shown in b expressed as relative binding, which refers to the DNA-binding activity of the HD-Zip-2 domain.

augment the DNA-binding specificity of a homeodomain.

It is intriguing that the evolution of the HD-Zip protein family occurred only in plants. Plant development is distinguished by its plasticity, which arises from a close coupling of developmental responses to environmental stimuli. Several pieces of evidence suggest that HD-Zip proteins may control developmental pathways peculiar to plants, such as growth responses to environmental signals (Carabelli *et al.*, 1993, 1996; Schena *et al.*, 1993; Baima *et al.*, 1995; Soderman *et al.*, 1996). In view of this fact, the ability of the HD-Zip proteins to bind DNA only in the dimeric form may confer specific advantages in regulating plant developmental processes. HD-Zip proteins might be turned into specific DNA-binding factors simply by promoting the formation of homodimeric complexes, and they might rapidly induce the changes in gene expression that underlie plant growth responses to distinct environmental stimuli.

Acknowledgements

We thank Roberto Gargamelli for photographic assistance. This work was supported partly by the Fondazione Istituto Pasteur-Fondazione Cenci Bolognetti, Università di Roma La Sapienza, and by Piano Nazionale 'Biotecnologie Vegetali', Ministero per il Coordinamento delle Politiche Agricole, Alimentari e Forestali.

References

Baima, S., Nobili, F., Sessa, G., Lucchetti, S., Ruberti, I. & Morelli, G. (1995). The expression of the *ATHB-8* homeobox gene is restricted to provascular cells in *Arabidopsis thaliana*. *Development*, **121**, 4171–4182.

Blackwell, T. K., Bowerman, B., Priess, J. R. & Weitraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZip and homeodomain elements. *Science*, **266**, 621–628.

Carabelli, M., Sessa, G., Baima, S., Morelli, G. & Ruberti, I. (1993). The *Arabidopsis ATHB-2* and *-4* genes are strongly induced by far-red-rich light. *Plant J.* **4**, 469–479.

Carabelli, M., Morelli, G., Whitelam, G. & Ruberti, I. (1996). Twilight-zone and canopy shade induction of the *ATHB-2* homeobox gene in green plants. *Proc. Natl Acad. Sci. USA*, **93**, 3530–3535.

Di Cristina, M., Sessa, G., Dolan, L., Linstead, P., Baima, S., Ruberti, I. & Morelli, G. (1996). The *Arabidopsis* ATHB-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J.* **10**, 393–402.

Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted a helices: crystal structure of the protein-DNA complex. *Cell*, **71**, 1223–1237.

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. & Wuthrich, K. (1994). Homeodomain-DNA recognition. *Cell*, **78**, 211–223.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61–68.

Keller, W., Konig, P. & Richmond, T. J. (1995). Crystal structure of a b-Zip/DNA complex at 2.2 Å: determinants of DNA specific recognition. *J. Mol. Biol.* **254**, 657–667.

Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579–590.

Konig, P. & Richmond, T. J. (1993). The X-ray structure of the GCN4-b-Zip bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. *J. Mol. Biol.* **233**, 139–154.

Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995). Crystal structure of the MATa1/MATα2 homeodomain heterodimer bound to DNA. *Science*, **270**, 262–269.

Mattsson, J., Soderman, E., Svenson, M., Borkird, C. & Engstrom, P. (1992). A new homeobox-leucine zipper gene from *Arabidopsis thaliana*. *Plant Mol. Biol.* **18**, 1019–1022.

Meijer, A. H., Scarpella, E., van Dijk, E. L., Qin, L., Taal, A. J. C., Rueb, S., Harrington, S. E., McCouch, S. R., Schilperoort, R. A. & Hoge, J. H. C. (1997). Transcriptional repression by Oshox1, a novel homeodomain leucine zipper protein from rice. *Plant J.* **11**, 263–276.

Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J. & Wuthrich, K. (1990). Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* **9**, 3085–3092.

Ruberti, I., Sessa, G., Lucchetti, S. & Morelli, G. (1991). A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. *EMBO J.* **10**, 1787–1791.

Sarker, G. & Sommer, S. S. (1990). The megaprimer method of site-directed mutagenesis. *Biotechniques*, **8**, 404–407.

Schena, M. & Davis, R. W. (1992). HD-Zip proteins: members of an *Arabidopsis* homeodomain protein superfamily. *Proc. Natl Acad. Sci. USA*, **89**, 3894–3898.

Schena, M., Lloyd, A. M. & Davis, R. W. (1993). The *HAT4* gene of *Arabidopsis* encodes a developmental regulator. *Genes Dev.* **7**, 367–379.

Sessa, G., Morelli, G. & Ruberti, I. (1993). The ATHB-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. *EMBO J.* **12**, 3507–3517.

Sessa, G., Carabelli, M., Ruberti, I., Lucchetti, S., Baima, S. & Morelli, G. (1994). Identification of distinct families of HD-Zip proteins in *Arabidopsis thaliana*. In *NATO ASI Series, Volume H81. Plant Molecular Biology* (Coruzzi, G. & Puigdomenech, P., eds), pp. 412–426, Springer Verlag, Berlin.

Soderman, E., Mattsson, J. & Engstrom, P. (1996). The *Arabidopsis* homeobox gene *ATHB-7* is induced by water deficit and by abscisic acid. *Plant J.* **10**, 275–281.

Suckow, M., Schwamborn, K., Kisters-Woike, B., von Wilcken-Bergmann, B. & Muller-Hill, B. (1994). Replacement of invariant bZip residues within the basic region of the yeast transcriptional activator GCN4 can change its DNA binding specificity. *Nucl. Acids Res.* **22**, 4395–4404.

Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991). Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*, **67**, 517–528.

Edited by T. Richmond

(Received 21 April 1997; received in revised form 20 August 1997; accepted 5 September 1997)